Midterm 3 note sheet

Table 1: Discriminant analysis.

Generative classifier prediction	$\arg\max_k \pi_k f_k(\underset{\sim}{x})$
Class prior	π_k
Class density	$f_k(\underset{\sim}{x})$
Class posterior	$\Pr(Y = k \mid \underset{\sim}{x}) = \frac{\pi_k f_k(\underset{\sim}{x})}{\sum_{l=1}^{K} \pi_l f_l(\underset{\sim}{x})}$
Normal density of one feature	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$
LDA model	$\underset{\sim}{x} \mid Y = k \sim \mathrm{N}(\vec{\mu}_k, \Sigma)$
Linear discriminant function	$\delta_k(\underset{\sim}{x}) = \underset{\sim}{x}^\top (\Sigma)^{-1} \vec{\mu}_k - \frac{1}{2} \vec{\mu}_k^\top (\Sigma)^{-1} \vec{\mu}_k + \ln \pi_k$
QDA model	$\underset{\sim}{x} \mid Y = k \sim \mathrm{N}(\vec{\mu}_k, \Sigma_k)$
Quadratic discriminant function	$\delta_k(\underset{\sim}{x}) = -\frac{1}{2} \ln\lvert\Sigma_k\rvert - \frac{1}{2} (\underset{\sim}{x} - \vec{\mu}_k)^\top (\Sigma_k)^{-1} (\underset{\sim}{x} - \vec{\mu}_k) + \ln \pi_k$
Discriminant prediction	$\arg\max_k \delta_k(\underset{\sim}{x})$
Two-class LDA posterior	$\Pr(Y = 1 \mid \underset{\sim}{x}) = \operatorname{expit}\left(\beta_0 + \vec{\beta} \cdot \underset{\sim}{x}\right)$
Its slope vector	$\vec{\beta} = (\Sigma)^{-1} (\vec{\mu}_1 - \vec{\mu}_0)$
Its intercept	$\beta_0 = \ln \frac{\pi_1}{\pi_0} - \frac{1}{2} (\vec{\mu}_1 + \vec{\mu}_0)^\top \vec{\beta}$
Logistic function	$\operatorname{expit}(z) = \frac{1}{1+\mathrm{e}^{-z}}$
Probability to log odds	$\pi \underset{\omega/(1+\omega)}{\overset{\pi/(1-\pi)}{\rightleftarrows}} \omega \underset{\mathrm{e}^\eta}{\overset{\ln \omega}{\rightleftarrows}} \eta$
Estimated prior	$\hat{\pi}_k = n_k / n$
Estimated class mean	$\hat{\underset{\sim}{\mu}}_k = \frac{1}{n_k} \sum_{i:y_i=k} \underset{\sim}{x}_i$
Estimated shared covariance	$\hat{\Sigma} = \frac{1}{n-K} \sum_{k=1}^{K} \sum_{i:y_i=k} (\underset{\sim}{x}_i - \hat{\underset{\sim}{\mu}}_k)(\underset{\sim}{x}_i - \hat{\underset{\sim}{\mu}}_k)^\top$

Table 2: Model selection.

Mean squared error	$\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$
K-fold cross-validation estimate	$\frac{1}{K} \sum_{k=1}^{K} \mathrm{error}_k$
Error of one fit	error_k = error on fold k of the fit to the other $K-1$ folds
Training points per fit	$\frac{K-1}{K} n$
Leave-one-out cross-validation	$K = n$

Table 3: Regularization.

Ridge regression objective	$\left\lVert X\vec{\beta} - \underset{\sim}{y}\right\rVert^2 + \lambda \left\lVert\vec{\beta}\right\rVert^2$
Ridge minimizer	$\vec{\beta}_{\text{ridge}} = (X^\top X + \lambda I)^{-1} X^\top \underset{\sim}{y}$
Lasso objective	$\left\lVert X\vec{\beta} - \underset{\sim}{y}\right\rVert^2 + \lambda \left\lVert\vec{\beta}\right\rVert_1$
Lasso penalty norm	$\left\lVert\vec{\beta}\right\rVert_1 = \sum_j \lvert\beta_j\rvert$
Soft-thresholding operator	$S_\tau(z) = \operatorname{sign}(z) \max(\lvert z\rvert - \tau, 0)$
Orthonormal features	$X^\top X = I$
Least squares coefficient there	$\hat{\beta}_j = X_{\cdot j} \cdot \underset{\sim}{y}$
Ridge coefficient there	$\frac{\hat{\beta}_j}{1+\lambda}$
Lasso coefficient there	$S_{\lambda/2}\left(\hat{\beta}_j\right)$
Inverse of a 2-by-2 matrix	$\begin{pmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\end{pmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

Table 4: Logarithms and exponentials.

Arguments of the logarithm	$a, b > 0$
Log of one	$\ln 1 = 0$
Exponential of zero	$\mathrm{e}^0 = 1$
Log of an exponential	$\ln(\mathrm{e}^s) = s$
Exponential of a log	$\mathrm{e}^{\ln a} = a$
Log of a reciprocal	$\ln(1/a) = -\ln a$
Exponential of a negative	$\mathrm{e}^{-s} = 1/\mathrm{e}^s$
Log of a product	$\ln(ab) = \ln a + \ln b$
Log of a quotient	$\ln(a/b) = \ln a - \ln b$
Log of a power	$\ln(a^k) = k \ln a$
Exponential of a sum	$\mathrm{e}^{s+t} = \mathrm{e}^s \mathrm{e}^t$
Exponential of a difference	$\mathrm{e}^{s-t} = \mathrm{e}^s / \mathrm{e}^t$
Power of an exponential	$(\mathrm{e}^s)^k = \mathrm{e}^{ks}$

Table 5: Decimal values of exponentials and logarithms.

	e	e^{-1}	e^{-2}	$\ln 2$	$\ln 3$	$\ln 4$	$\ln 5$	$\ln 10$
Value	2.718	0.368	0.135	0.693	1.099	1.386	1.609	2.303